# CS7643: Deep Learning
# Fall 2017
# HW0 Solutions

Dong-Hyun (Tony) Kim

February 22, 2020

## 1 Probability and Statistics

1. No.

$$\begin{aligned}\mathbb{E}[x] &= \sum_x x * p(x) \\ &= 1 * 1/6 + \sum_2^6 (-1/4) * 1/6 \\ &= 1/6 - 5/6 \\ &= -4/6\end{aligned}$$

2. The corresponding distributive function is

$$\begin{cases} \int_0^x 4x dx = 2x^2 \Big|_0^x & 0 \leq x \leq 1/2 \\ \frac{1}{2} + \int_{\frac{1}{2}}^x -4x + 4dx = \frac{1}{2} + -2x^2 + 4x \Big|_{\frac{1}{2}}^x & \frac{1}{2} \leq x \leq 1 \\ 1 & 1 \leq x \end{cases}$$

3.

$$\begin{aligned}\mathbb{E}[x - \mathbb{E}[x]]^2 &= \mathbb{E}[x^2 - 2x\,\mathbb{E}[x] + \mathbb{E}[x]^2] \\ &= \mathbb{E}[x^2] - 2\,\mathbb{E}[x]\,\mathbb{E}[x] + \mathbb{E}[x]^2 \\ &= \mathbb{E}[x^2] - 2\,\mathbb{E}[x]^2 + \mathbb{E}[x]^2 \\ &= \mathbb{E}[x^2] - \mathbb{E}[x]^2\end{aligned}$$

4.

$$\begin{aligned}\int_{-\infty}^{\infty} p(x)(ax2 + bx + c)dx &= \mathbb{E}[(ax2 + bx + c)] \\ &= a\,\mathbb{E}[x^2] + b\,\mathbb{E}[x] + c \\ &= a * (Var(x) - \mathbb{E}[x]^2) + b * 0 + c \\ &= a * (1 - 0) + 0 + c \\ &= a + c\end{aligned}$$

Proving Stuff
5.

5) $\log_e x \le x - 1$

$$\log_e x - (x-1) \le 0$$

differentiate

$$\frac{1}{x} - 1 \le 0$$

$$\frac{1}{x} - 1 = 0 \iff x = 1$$

Second derivative

$$\frac{d}{dx} \cdot x^{-1} - 1 \le 0$$

$$-x^{-2} \le 0$$

6) $KL(p, q) = \sum_{i=1}^{K} p_i \log\left(\frac{p_i}{q_i}\right)$

Motivation: KL divergence is a good metric to measure performance for ML models. We show why through this problem.

a) $KL(p, q)$ is always positive

$$\sum_{i=1}^{K} p_i \log\left(\frac{p_i}{q_i}\right) = -\sum_{i=1}^{K} p_i \log\left(\frac{q_i}{p_i}\right) \cancel{\geq} -\sum_{i=1}^{K} p_i \left(\frac{q_i}{p_i} - 1\right)$$

$$= -\sum_{i=1}^{K} q_i - p_i$$

$$= \sum_{i=1}^{K} p_i - \sum_{i=1}^{K} q_i$$

$$= 0$$

b) $\sum_{i=1}^{K} p_i \log\left(\frac{p_i}{q_i}\right) = \sum_{i=1}^{K} p_i (\log p_i - \log q_i) = 0$

$\downarrow$ differentiate

$$\sum_{i=1}^{K} \frac{p_i}{p_i} - \sum_{i=1}^{K} \frac{p_i}{q_i} = 0$$

$$\sum_{i=1}^{K} \frac{p_i}{p_i} = \sum_{i=1}^{K} \frac{p_i}{q_i}$$

$$\boxed{p = q}$$

c) Not symmetric when $p = \begin{cases} \frac{1}{3}, & x=1 \\ \frac{1}{3}, & x=2 \\ \frac{1}{3}, & x=3 \end{cases}$ and $q = \begin{cases} \frac{2}{3}, & x=1 \\ \frac{1}{6}, & x=2 \\ \frac{1}{6}, & x=3 \end{cases}$

Counter example.

Math 3012

Homework

Dong-Hyun (Tony) Kim

1. $\frac{6 \cdot 6}{2} = \boxed{18}$ 6 of them are reflexive

2. reflexive

Suppose $a \bmod m = n$. Then, $a \bmod m = n = a \bmod m$

Now, for symmetry, Suppose $a \bmod m = n = b \bmod m$.

$b \bmod m = n = a \bmod m$.

Lastly, transitivity. Suppose $a \bmod m = n = b \bmod m$ and $b \bmod m = n = c \bmod m$. Then, $a \bmod m = n = c \bmod m$

5. $x_1$ : $\{1, 3, 4, 5, 6\}$

$x_2$ : $\{1, 2, 4, 5, 6\}$

$x_3$ : $\{1, 2, 3, 6\}$

$x_4$ : $\{1, 2, 3\}$

$x_5$ : $\{1, 5, 6\}$

$x_6$ : $\{1, 3, 6\}$

$x_7$ : $\{1, 2\}$

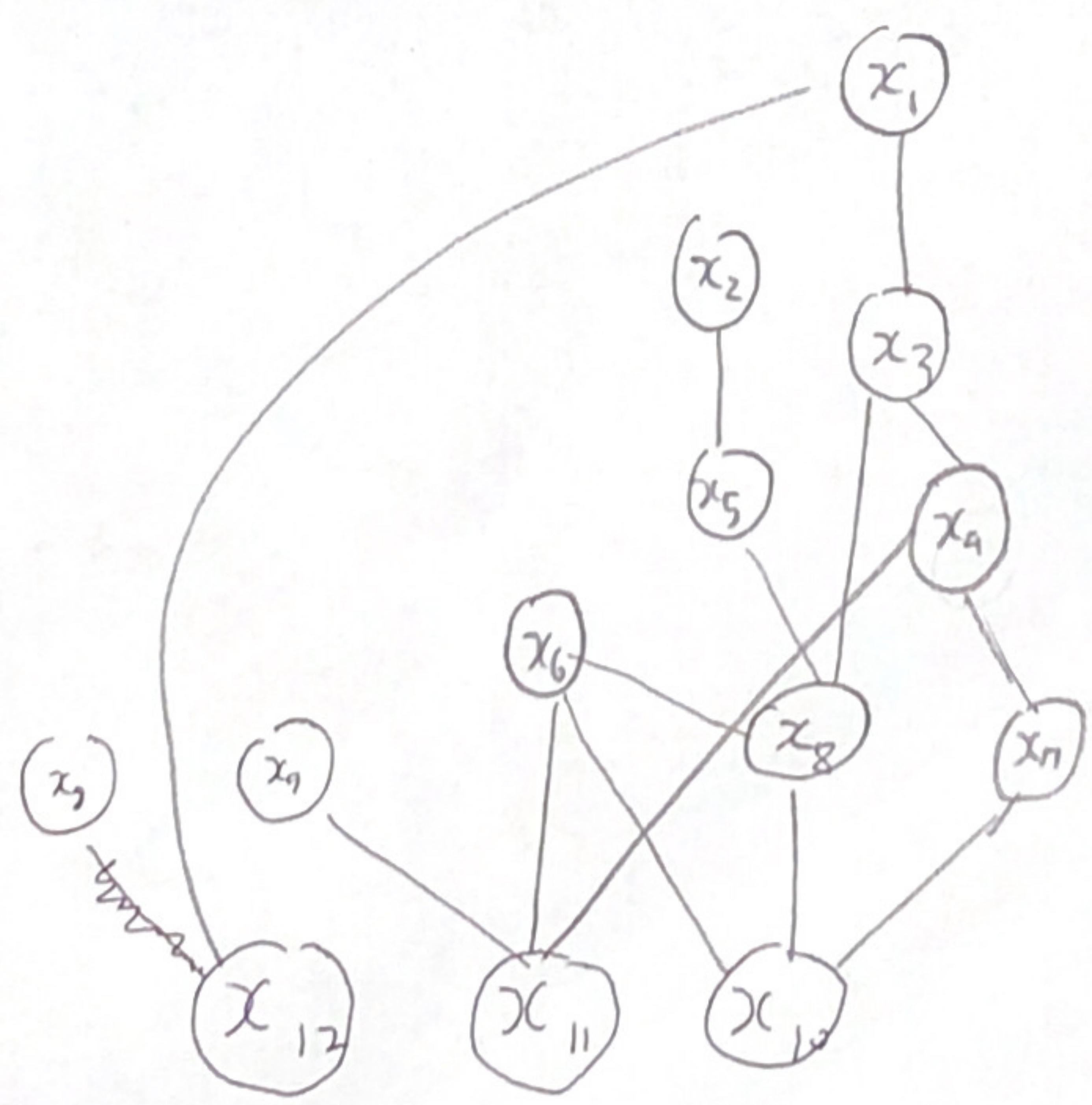
$x_8$ : $\{1, 6\}$

$x_9$ : $\{3, 5\}$

$x_{10}$ : $\{1\}$

$x_{11}$ : $\{3\}$

$x_{12}$ : $\{4\}$



8.

a) 15, 8, 11, 2, 17, 3

b) 16, 1, 5, 14

c) 14 → 2

d) 14, 9, 3 , {14, 9, 6, 10, 2} is bigger.

e) 16, 1, 5, 14

9

$A_1$ : {22, 18, 23, 12}

$A_2$ : {13, 3, 2, 21, 11, 17}

$A_3$ : {1, 5, 25, 24, 10}

$A_4$ : {8, 20}

$A_5$ {9, 19, 16}

$A_6$ {6, 7}

$A_7$ {14, 26}

$A_8$ {15}

height: 8

maximum chain: {22, 3, 4, 29, 9, 7, 26, 15}

11) According to Sperner's Theorem,

$$\binom{10}{5} = \boxed{\frac{10!}{5!5!}}$$